

# 3차원 비디오 전송을 위한 스테레오비디오 동기화 방법

정회원 이 동 진\*, 이 선 오\*, 종신회원 심 동 규\*, 이 혁 준\*\*

## Stereo-video Synchronization for 3D Video Transmission

Dong-Jin Lee\*, Seon-Oh Lee\* *Regular Members,*  
Dong-Gyu Sim\*, Hyuk-Joon Lee\*\* *Lifelong Members*

### 요 약

본 논문에서는 3차원 효과의 최대화를 위한 스테레오 비디오 전송 방법을 제안한다. 기존의 다중스트림 동기화 기술은 단일 비디오와 오디오간의 지연을 최소화 하는 것으로, 둘 이상의 비디오가 동시에 필요한 3차원 비디오 스트림 동기화에는 적합하지 않다. 본 논문에서는 인간에게 자연스럽게, 입체적으로 느껴지는 3차원 영상을 만들어 낼 수 있는 시간 차이의 허용 범위를 고려하여 다중 영상을 합성하는 기준을 제시하였고, 이를 위한 비디오 스트림 간의 동기화 방법을 제안하였다. 본 연구에서는 주관적 화질평가를 통하여 스테레오 비디오를 3차원으로 느낄 수 있는 시간 차이의 허용범위를 측정하였다. 스테레오 비디오의 동기화를 위해, 전송된 각 비디오 스트림들의 취득 시간을 구하고, 시간 차이 허용 범위 안에 포함된 비디오 스트림을 이용하여 3차원 영상으로 합성한 뒤, 디스플레이 장치를 통해 재생한다. 제안한 기술의 성능을 평가하기 위해 실시간 다중 비디오 통신 시스템을 구현하였고, 주관적 화질 평가를 통해 제안한 동기화 제어 기술의 성능을 평가하였다. MOS (Mean Opinion Score) 측정 결과, 제안하는 기술을 통하여 3차원 디스플레이 장치에 재생한 영상은 DMOS (Differential Mean Opinion Score) 실험 스케일 중 매우 좋음과 좋음 범위에 속하는 것을 확인하였다.

**Key Words** : Stereo-video, multi-view, RTP/RTCP, synchronization, 3D

### ABSTRACT

In this paper, we propose a stereo-video transmission method for reduction of delay and maximization of 3D effect. Conventional multimedia synchronization algorithms were designed to achieve minimum delay and synchronize multiple video and audio streams, however, they could not be effective for 3D video transmission. In this paper, we proposed a synchronization algorithm by considering the minimum error of time difference between streams for 3D effect. The minimum error of time difference for 3D effect was derived based on a 3D subjective quality test. We compute display time of the delivered videos within the allowed time-difference and the video are displayed according to the display time. To evaluate the performance of the proposed algorithm, we implemented a real-time video communication system and subjective quality test has been conducted with the proposed system. We found that video quality displayed by the proposed system. We found that video quality displayed by the proposed algorithm ranks 'good' and 'excellent' in the DMOS (Differential Mean Opinion Score) scale, based on the MOS (Mean Opinion Score) test.

\* 이 논문은 2008년도 광운대학교 교내학술연구비 지원에 의해 연구되었음

\* 광운대학교 컴퓨터공학과 영상처리시스템연구실 (dgsim@kw.ac.kr)

\*\* 광운대학교 컴퓨터공학과 이동컴퓨팅연구실 (hlee@kw.ac.kr)

논문번호 : KICS2008-11-492, 접수일자 : 2008년 11월 3일, 최종논문접수일자 : 2009년 3월 24일

## I. 서 론

패킷 기반 네트워크의 대역폭 증가로 인하여, 실시간으로 고품질 영상을 전송하는 새로운 멀티미디어 서비스 통신 시스템 구현이 가능해졌다. 이에 IPTV, 화상 회의 서비스 등이 활발하게 사용되고 있다. 이와 같은 서비스의 고품질화와 함께, 최근 들어 3차원 비디오에 대한 관심과 요구가 증가되고 있다.

3차원 비디오에 대한 연구는 주로 3차원 영상을 취득하는 방법과 이를 디스플레이 하는 방법에 대한 연구를 중심으로 진행되어 왔다. 3차원 영상은 한 대 이상의 카메라를 통해 동 시간대에 다양한 시점에서 촬영된 영상들로 만들어지는데, 이를 위한 카메라의 위치 및 각도, 필요한 카메라의 수, 각 카메라에서 취득한 영상을 저장하는 방법 등 여러 가지 연구가 진행되어 왔다. 취득된 영상들로 3차원 영상을 만들어 내는 것도 중요하지만, 이를 사람에게 3차원으로 인식하게 만드는 기술에 대한 연구 역시 중요하고, 이는 비교적 오래 진행되어 왔다. 우리가 물체를 3차원으로 인식할 수 있는 것은 하나의 물체를 바라볼 때 오른쪽 눈과 왼쪽 눈으로 들어오는 신호가 다르기 때문이다. 이를 근거로 양쪽 눈으로 들어가는 영상 신호를 다르게 함으로써, 사람으로 하여금 입체감을 느낄 수 있게 한다. 이를 위해 편광 안경을 이용하거나, 디스플레이 되는 장치에 렌즈를 부착함으로써 각 눈으로 들어가는 영상이 다르게 만드는 방법 등으로 연구되어 왔다<sup>[1]</sup>. 최근에는 영상의 취득과 디스플레이과정 사이에 필요한 기술에 대한 연구가 활발하게 진행되고 있다. 이는 비디오 코딩 기술과 관련된 표준화를 주로 진행해오던 MPEG (Moving Pictures Experts Group) 과 JVT (Joint Video Team)에서 연구 및 표준화를 주도하고 있다<sup>[2]</sup>. 이 두 단체 역시 처음에는 부호화 효율과 관련된 기술을 위주로 연구 및 표준화를 진행해 왔으나, 최근 표준화 과정에서 서비스 실용화에 필요한 다양한 기술들이 제안되면서, 이와 관련된 표준화도 동시에 진행하고 있다. 3차원 서비스에 필요한 연구가 위의 세 가지로 진행되어 왔으나, 실제 이러한 서비스를 진행하기 위해서 필요한 시스템 동기화 기술은 아직까지 거의 진행되지 않고 있다. 패킷을 통해 영상을 전송하게 되면, 같은 시점의 여러 영상을 동시에 보낼 수 없기 때문에 이를 전송받는 수신부에서 전송된 각 영상을 동기화 시켜주는 기술이 절대적으로 필요하다. 본 논문에서는 스테레오 비디오를 위한 다중 비디오 스트림 동기

화 방법을 제안한다. 멀티미디어 시스템 상에서 발생한 지연들을 가진 스트림들을 동기화하는 방법들은 3차원 비디오를 고려하여 설계되어 있지 않다. 기존의 다중 비디오 방법은 3차원 비디오를 위해서라기보다는 다중 비디오, 오디오, 및 메타 데이터와 같은 다중 스트림을 동기하기 위한 방법이었다. 이러한 방법들은 여러 개의 스트림을 동기화시키기 위해서 많은 지연이 요구하는 단점을 가지고 있다. 본 논문에서는 3차원 디스플레이 장치를 통해 재생했을 때, 사람들에게 자연스럽게 입체적으로 느껴지는 지연 시간의 허용범위를 측정하고, 그 범위를 이용하여 비디오스트림 간의 훼손된 시간 관계를 자연스럽게 복원하는 방법에 기반을 두고 있다. 또한, 지터가 적은 비디오 재생을 위하여, 네트워크에서 발생하는 비디오 스트림의 손실 보정 방법을 제안하였다. 본 논문에서는 제안한 동기화 제어 기술의 성능을 평가하기 위해 실시간 다중 비디오 통신 시스템을 구현하였고, 지연 및 지터 등이 포함된 네트워크 환경에서 스테레오 영상을 송신 및 수신하여, 제안한 기술의 유효성을 검증하는 실험을 하였다. 스테레오 영상은 동시에 두 개의 영상만이 존재하기 때문에 영상간의 시간차이 범위를 정확하게 측정할 수 있다. 또한, 구현한 실시간 다중 비디오 통신 시스템을 통해 제어된 3차원 영상을 이용하여 DMOS 실험을 통해 주관적 화질 평가를 수행하였다.

본 논문의 2 장에서는 3차원 영상의 주관적 화질 측정 방법에 대하여 기술하였다. 3 장에서는 기존의 RTP/RTCP를 이용한 실시간 멀티미디어 통신 시스템에서의 동기화 기술에 대하여 설명하고, 4 장에서는 제안한 다중 비디오 스트림 간의 동기화 제어 기술에 대해 설명 한다. 5 장에서는 제안한 기술을 다중 비디오 통신 시스템에 적용한 실험한 결과를 제시하고, 이를 분석한다. 마지막으로, 6 장에서는 결론을 제시한다.

## II. 3차원 영상의 주관적 화질 측정

스테레오 영상을 디스플레이 하는데 있어서 가장 중요한 것은 동시에 뿌려지는 양쪽 뷰 간의 목표 지연 시간과의 차이를 어느 정도까지 허용 하는가의 여부이다. 뷰 간 목표 지연 시간과의 오차 한계를 알아보기 위하여, 오차를 여러 가지로 설정한 영상들로 화질 평가를 수행하였다.

영상의 화질을 평가하는 방법은 크게 두 가지로, 사람이 직접 영상을 보고 평가하는 주관적 화질 측

정 방법과 수학적 모델을 이용한 화질 평가 알고리즘을 통해 평가하는 객관적 화질 측정 방법이 있다. 주관적 화질 평가 방법은 사람이 영상을 보고 자신이 느끼는 화질의 정도를 주관적으로 평가하는 것으로, 객관적 화질 측정에 의한 평가 결과보다 유효하고 신뢰도 높은 결과를 도출한다. 주관적 화질 평가 절차에 대한 ITU-R의 표준 권고안 BT.500-11에서 정의된 주관적 화질 측정 방법은 2차원 영상에 한해서 적용되는 방법이지만, 아직까지 3차원 영상에 대한 주관적 화질 측정 방법이 정의 되어 있지 않은 상태이고, 주관적 화질 측정 절차에 대한 방법에서는 사용 영상 자체에 대해 제한을 두지 않기 때문에, 해당 권고안의 표준 절차를 적용함에 있어 문제가 없을 것으로 판단되어 이를 적용하였다<sup>3)</sup>. 여기에는 DSCQS (Double-stimulus continuous quality-scale), DSIS (Double-stimulus impairment scale), SDSCE (Simultaneous double stimulus for continuous evaluation), ACR (Absolute category rating) 등 여러 가지 방법이 있는데, 본 논문에서는 ITU-R BT.500-11 권고안의 방법 중 하나인 ACR 방법을 사용한다. 이는 여러 조건의 영상을 무작위로 섞어 한 번씩만 보여주기 때문에, 각 조건의 영상마다 느끼는 주관적 화질간의 의존성이 적기 때문에 오차 한계를 알아보기 위한 본 실험에서는 가장 적합한 방법이고, 다른 방법에 비해 테스트에 필요한 시간이 적은 장점도 있어 해당 방법을 선택하였다.

ACR 방법에서, 평가자는 각 조건의 영상을 보고 이에 대한 품질을 5가지 (매우 좋음; Excellent(5), 좋음; Good(4), 보통; Fair(3), 떨어짐; Poor(2), 나쁨; Bad(1)) 중 하나로 평가한다. 각 조건의 영상에 대한 사람들의 주관적인 화질 측정값들을 이용하여 정규화시켜 얻을 수 있는 값을 MOS라 한다. 이 값은 다음과 같이 계산된다.

$$MOS_i = \frac{1}{n} \sum_{j=1}^n m_{ij} \quad (1)$$

MOS<sub>i</sub> 는 i 번째 영상에 대한 평균 화질 측정값이고, n는 평가자의 총 수, 그리고 m<sub>ij</sub>는 i 번째 영상에 대해서 j 번째 평가자가 측정하여 정규화 시킨 값이다.

본 논문에서는 평가 후, 같은 영상에 대해 시간 오차가 없는 원본 조건과 시간 오차가 있는 조건의 결과 쌍을 모아 두 조건의 평가 결과 차이를 얻는다. 이 차이값의 평균을 통해 여러 사람에 의한 화질 평

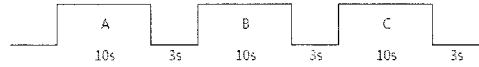


그림 1. 시간에 따른 ACR 테스트 비디오 재생

가 결과를 하나의 값으로 만드는데 이를 DMOS라 한다. 이는 평가자들이 부여한 동영상에 대한 점수 차이의 평균을 의미한다. 즉 영상간의 시간차가 존재하지 않는 원본영상의 MOS 값 (MOS<sub>r</sub>)에서 영상간의 시간차이가 발생한 처리영상의 MOS 값 (MOS<sub>p</sub>)의 차이가 처리 영상의 DMOS 값이 된다.

$$DMOS = MOS_r - MOS_p \quad (2)$$

이 차이들 중 주관적으로 허용 가능한 범위에 있는 조건들을 분석하여, 영상처리에 사용 가능한 시간차이 정도를 파악한다. 그림 1은 ACR 테스트 방법에 따른 비디오의 재생 과정에 대해서 보여주고 있다.

3차원 영상을 주관적으로 평가하기 위해서 3차원 디스플레이 장치인 DIMEN C170S를 이용하여 1024×768 해상도에서 320×240 크기의 합성한 3차원 영상을 보여준다. 3차원 영상에 대한 주관적인 화질 측정에 대한 신뢰성을 높이기 위해, 평가자가 한 명씩 화질 평가에 참여했다. 비디오 프레임 간의 시간차이를 효율적으로 분석하기 위하여, 다중 비디오로 스테레오 영상을 사용하였다. 스테레오 영상의 경우는 동시에 2개의 영상만이 존재하여 영상간의 시간차이 범위를 정확하게 측정할 수 있기 때문에 본 실험에 적합하다. 평가자들은 비전문가들로 구성된 15명이 참여했다. 이는 ITU-R BT.500-11 권고안에서 규정하고 있는 주관적 화질 측정을 위한 최소 인원 규정을 만족한다<sup>3)</sup>.

그림 2는 3차원 영상에 대한 DMOS 측정 결과를 나타내고 있다. 식 (5)에서와 같이 좌 우측의 프레임의 시간차이(단위:ms)를 변경한 영상 (processed)의 MOS 값 (MOS<sub>p</sub>)에서 동일한 시간에 취득한 영상 (reference)에 대해 측정한 MOS (MOS<sub>r</sub>)의 차이인 DMOS 값을 보여준다.

평가자들은 모든 영상에 대해서 DMOS 값이 2.5 이하 값을 가지는 스테레오 영상의 경우에 자연스럽게 입체적인 3차원 영상으로 인식하였다. ‘Mountain’ 영상의 경우는 테스트 영상의 움직임이 적기 때문에, 좌측과 우측 프레임 간의 시간차이가 큰 경우에도 사람들에게 자연스러운 3차원 영상으로 인식된다. 반면에 ‘Diving’, ‘Soccer2’와 ‘Trapeze’ 영상 등의 경우는 영상의 움직임이 빠르기 때문에 좌측

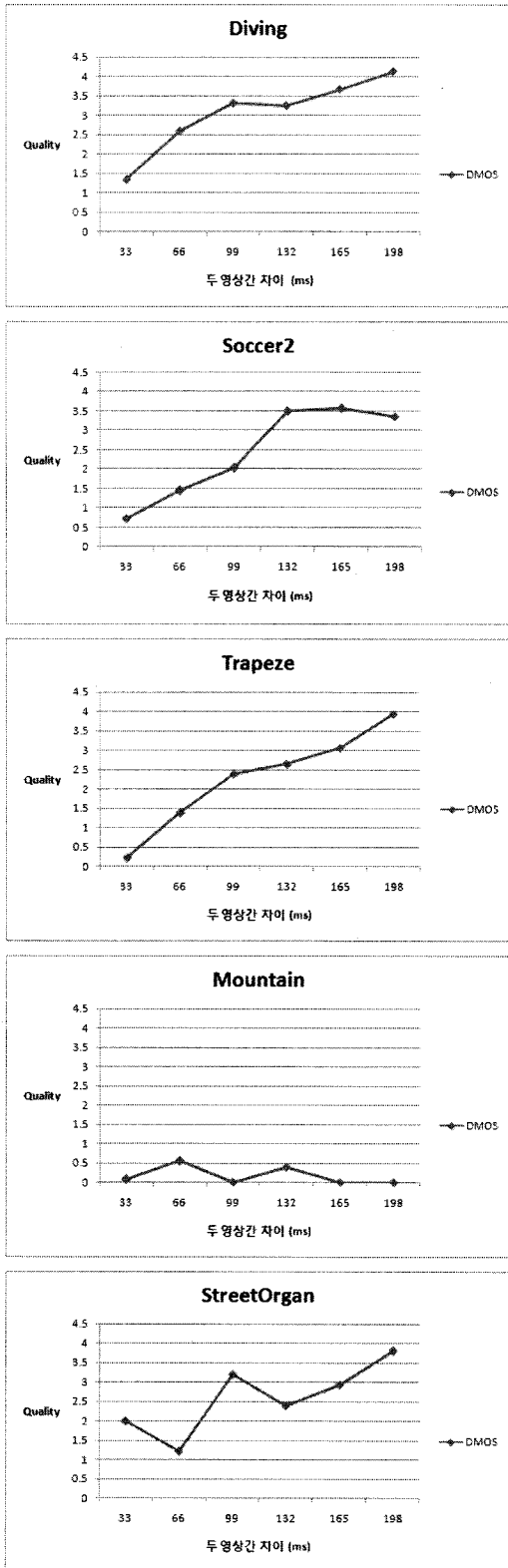


그림 2. 3차원 영상에 대한 DMOS 측정 값

과 우측 프레임간의 시간차이가 66ms 보다 큰 경우에는 3차원 영상으로는 자연스럽게 않다고 인식하는 경우가 많았다. ‘Street Organ’ 영상과 같은 경우처럼 몇몇 영상은 좌측과 우측 프레임의 시간 차이가 일정 정도 (예 - 99ms) 이상의 경우에는 사람이 3차원 영상으로 간간히 인식하는 경우도 발생하기도 한다.

모든 영상에 대해서 DMOS 값이 사람들이 3차원 영상으로 인식하는 범위인 2.5 이하를 찾아보면, 대부분의 영상에서 99ms 이하에서 발생하고 있다. 하지만, 그림 2의 ‘Diving’의 경우에 좌측과 우측의 프레임의 차이가 66ms일 때에 2.5 이하의 DMOS 값을 가진다. 우리가 측정을 하고자 하는 것은 3차원 영상으로 인식되는 좌측과 우측 프레임 간의 시간차이기 때문에, 최악의 경우까지 고려해야 한다. 따라서 3차원 영상을 구성하는 프레임 간의 시간차이가 66ms 경우라야 사람들이 자연스럽게 입체적인 3차원 영상으로 인식한다는 결론을 내릴 수 있다.

### Ⅲ. RTP/RTCP를 이용한 비디오 통신 시스템에서의 동기화

RTP (Real-time Transport Protocol)는 미디어의 타입에 따라 다양한 형태의 멀티미디어 데이터를 실시간으로 전송하도록 설계되어 있다. 본 장에서는 RTP/RTCP를 이용하여 실시간 멀티미디어 통신 시스템에서 동기화하는 기술을 살펴본다.

일반적으로 실시간 멀티미디어 통신 시스템에서 미디어 간의 동기화를 제어는 RTP/RTCP 패킷내의 RTP 타임스탬프와 NTP 타임스탬프를 이용함으로써 이루어진다. 이를 이용하여, 각 미디어 데이터가 취득된 시점에서 미디어간의 시간관계를 확인한다<sup>[4-6]</sup>. 그리고, 미디어 데이터들이 취득된 시점부터 재생하는 시점까지의 시간을 계산한다. 이렇게 미디어들이 재생까지 걸리는 시간들을 파악하여, 취득시의 시간 관계가 유지되고 있는 지를 확인하고, 만일, 취득한 시점의 시간관계가 훼손된 경우에는 미디어들이 재생될 때까지 걸리는 시간간의 차이 값을 구해서, 훼손된 미디어 간의 시간관계를 복원한다. 이러한 방식으로 미디어의 동기화를 제어 한다<sup>[7][8]</sup>.

오디오와 비디오 간의 동기화를 제어하는 기술에 대한 설명이 그림 3에 나타나 있다. 수신단에서 동기화를 제어하기 위해서는 그림 3에서와 같이 각각의 미디어에 해당하는 취득시간과 재생시간을 계산하여야 한다. 수신한 RTP 패킷 내의 미디어의 취득

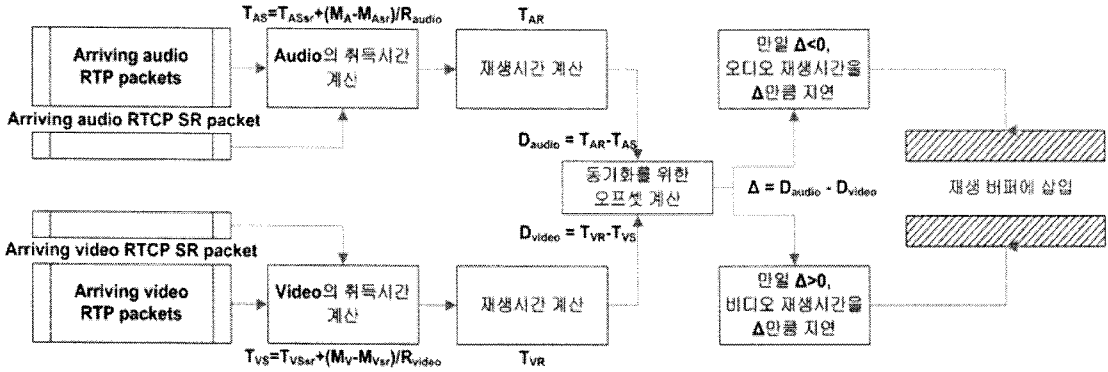


그림 3. 오디오와 비디오 간의 동기화 제어 기술

시간은 절대적인 시스템 시간이어야 하는데, RTP 패킷만으로는 절대적인 시스템 기준 시간을 알 수 없다. 왜냐하면 RTP 패킷 헤더에 있는 타임스탬프는 임의의 값으로 시작되어 미디어의 샘플링 비율에 따라 증가된 값이기 때문이다. 하지만, 수신한 RTCP 패킷의 NTP 타임스탬프가 있으면 RTP 패킷의 헤더에 있는 타임스탬프에 대한 절대적인 시스템 시간을 계산할 수가 있다. 따라서 미디어의 취득된 시간은 NTP 타임스탬프에 RTP 타임스탬프를 이용해서 계산하는데, 이 과정은 식 (3)에 나타나 있다<sup>9)</sup>.

$$T_s = T_{sr} + (M - M_{sr}) / Rate \quad (3)$$

$T_{sr}$ 은 RTCP 패킷에 있는 NTP 타임스탬프,  $M$ 은 RTP 패킷에 있는 RTP 타임스탬프,  $M_{sr}$ 은 RTCP 패킷에 있는 RTP 타임스탬프,  $Rate$ 는 미디어의 샘플링 레이트, 그리고  $T_s$ 는 미디어가 취득된 시간을 나타낸다. 미디어의 지연시간  $D$ 는 식 (4)를 통해, 식 (3)에서 구한 미디어의 취득시간  $T_s$ 과 미디어의 재생시간  $T_r$ 의 차이 값으로 계산된다. 여기서, 미디어의 지연시간에는 네트워크에서 생기는 지연시간과 미디어 코덱에서 발생할 지연시간이 포함되어 있다.

$$D = T_r - T_s \quad (4)$$

오디오와 비디오의 지연시간의 차이  $\Delta$ 는 위 식에서 구한 오디오의 지연시간 ( $D_{audio}$ )과 비디오의 지연시간 ( $D_{video}$ )의 차이 값이 된다.

$$\Delta = D_{audio} - D_{Video} \quad (5)$$

식 (5)을 통해서 오디오와 비디오의 지연시간의

차이  $\Delta$ 를 구할 수 있다. 이것은 오디오와 비디오의 네트워크와 시스템에서 발생한 지연시간의 차이가 된다. 즉, 두 미디어간의 취득시의 시간관계의 훼손 정도이다. 그래서 재생에서도 이러한 시간관계를 복원하기 위해서는 이 두 미디어간의 지연시간의 차이  $\Delta$ 를 이용해야 한다. 여기서,  $\Delta$ 의 값이 음수인 경우, 오디오 재생 시간을  $\Delta$ 값만큼 지연시킨 후 재생을 하고,  $\Delta$ 의 값이 양수인 경우에는 비디오 재생 시간을  $\Delta$ 값만큼 지연시킨 후 재생함으로써, 오디오와 비디오의 동기화를 제어 한다<sup>10-13)</sup>.

#### IV. 제안한 다중 비디오 스트림 제어 기술

본 장에서는 스테레오 입체 비디오를 위한 다중 비디오 스트림 간의 동기화 제어 기술을 소개한다. 자연스러운 3차원 영상으로 합성할 수 있는 다중 비디오 스트림 간의 시간차이 측정 결과를 분석하고, 이를 이용하여 다중 비디오 스트림 간의 동기화를 제어하는 기술에 대해서 설명한다.

일반적인 동기화의 개념은 해당 미디어를 구성하는 요소들이 취득된 시간 관계들을 유지하면서 재생되게 노력하는 것을 의미한다. 하지만, 다중 비디오 스트림을 재생하는 경우에는 취득한 다중 비디오 스트림을 각각의 디스플레이 장치에 출력하는 것이 아니라, 하나의 3차원 디스플레이 장치를 통하여 출력하게 된다. 이러한 다중 비디오 스트림 동기화 제어 기술은 하나 이상의 비디오 스트림 간의 시간 관계를 유지하는 것까지 고려해야 한다. 즉, 다중 비디오 스트림을 3차원 디스플레이 장치를 통해 재생했을 때, 사람들에게 자연스럽고 입체적인 3차원 영상으로 인식시키는 것까지 고려해야 한다는 것이다. 다중 비디오 스트림을 재생하는 경우에도

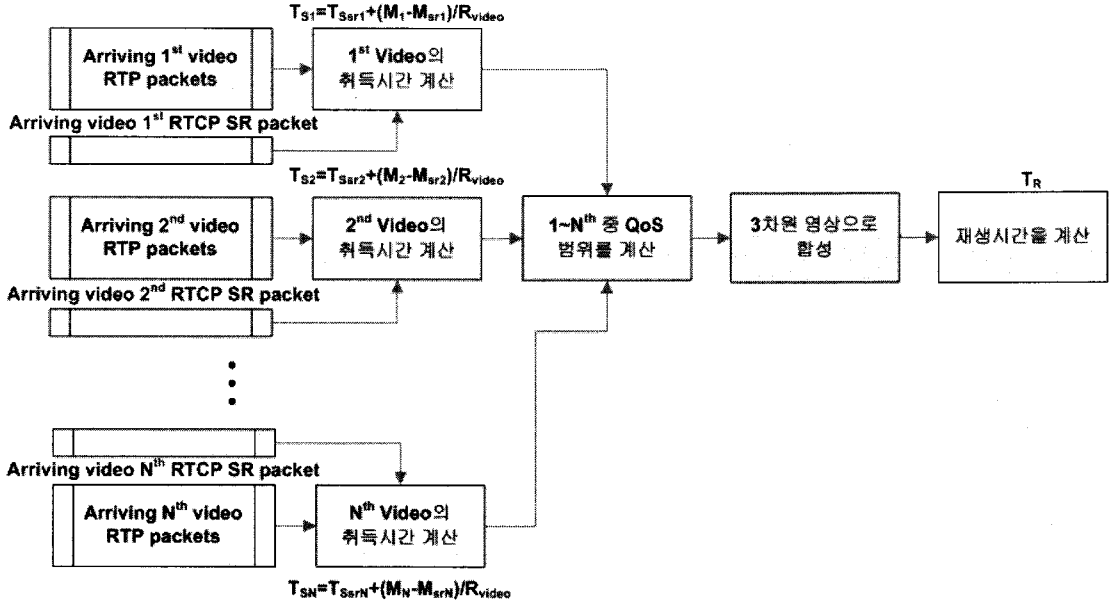


그림 4. 제안한 다중 비디오 스트림 간의 동기화

여러 종류의 스트림들을 합성하고 이를 사용자에게 보여줄 때 3차원 영상으로 느껴지는 시간차이 허용 범위를 측정해야 한다. 시간차이 허용범위는 2장에서 설명한 DMOS 실험을 통해 측정된 시간차이 허용범위 내에 취득된 다중 비디오 스트림을 합성하면, 사람들에게 자연스럽게 입체적으로 느껴지는 3차원 영상을 만들 수 있다. 이와 같은 사실을 기반으로 시간차이 허용범위 안에 포함된 비디오 스트림들을 이용하여 3차원 영상으로 합성하면, 비디오 스트림간의 동기화를 제어하는데 최소한의 지연을 통하여 자연스럽게 입체적인 3차원 영상을 얻을 수 있다<sup>[14]</sup>.

그림 4는 제안한 다중 비디오 스트림 간의 동기화를 제어하는 기술에 대한 블록도이다. 먼저, 네트워크를 통해 전송된 RTP/RTCP 패킷을 수신한다. 수신된 다중 비디오 스트림은 RTP/RTCP 패킷의 구문요소를 이용하여 해당 비디오의 취득 시간을 계산한다<sup>[15]</sup>. 취득 시간이 가장 빠른 비디오 스트림을 기준으로 자연스럽게 3차원 영상으로 합성되는 시간차이 허용범위 내에 존재하는 비디오 스트림이 있는 지를 확인한다. 비디오 스트림 중에 시간차이 허용범위를 벗어나는 것이 존재하면, 벗어난 비디오 스트림들은 3차원 영상을 합성할 때 사용하지 않고 나머지 비디오 스트림들을 이용하여 합성한다.

그림 5에는 시간차이 허용범위를 이용한 다중 비디오 스트림 간의 동기화 과정에 대해서 소개되어

있다. 이는 그림 4에서 수신된 다중 비디오 스트림 중에서 자연스럽게 입체적인 3차원 영상으로 합성되는 시간차이 허용범위 내에 있는 비디오 스트림을 찾아내는 부분을 자세히 설명하는 부분이다. 이 과정에서는 재생 예정인 다중 비디오 스트림들 중에서 가장 먼저 취득된 비디오 스트림을 확인한다. 다중 비디오 스트림 중에서 가장 먼저 취득된  $i$ 번째 비디오 스트림을 기준으로 시간차이 허용범위에 포함된 비디오 스트림들을 확인해서 가져온다. 이 시간차이 허용범위 내에 포함된 비디오 스트림들을 이용하여 3차원 영상으로 합성할 수 있다면, 시간차이 허용범위를 벗어나는 나머지 비디오 스트림은 사용하지 않는다. 만일 시간차이 허용범위 내에 자연스러운 3차원 영상으로 합성할 수 있는 비디오 스트림이 포함되어 있지 않으면, 사람들에게 덜 자연스럽게 3차원 영상으로 느껴지는 2차 시간차이 허용범위에 포함되는 비디오 스트림들을 확인해서 합성에 사용한다. 2차 시간차이 허용범위에서는 최근에 재생에 사용했던 비디오 스트림들이 허용범위에 포함된다면 사용한다. 2차 시간차이 허용범위에도 포함된 비디오 스트림이 없다면, 이전에 재생했던 3차원 영상을 연속해서 재생하도록 한다. 하지만, 이와 같이 이전의 3차원 영상을 그대로 재생하는 경우에는 영상의 움직임이 부드럽게 이어지지 않고 끊어져서 재생되는 현상이 발생할 수 있다. 이러한 현상을 최소화하기 위해서는 이전 3차원 영

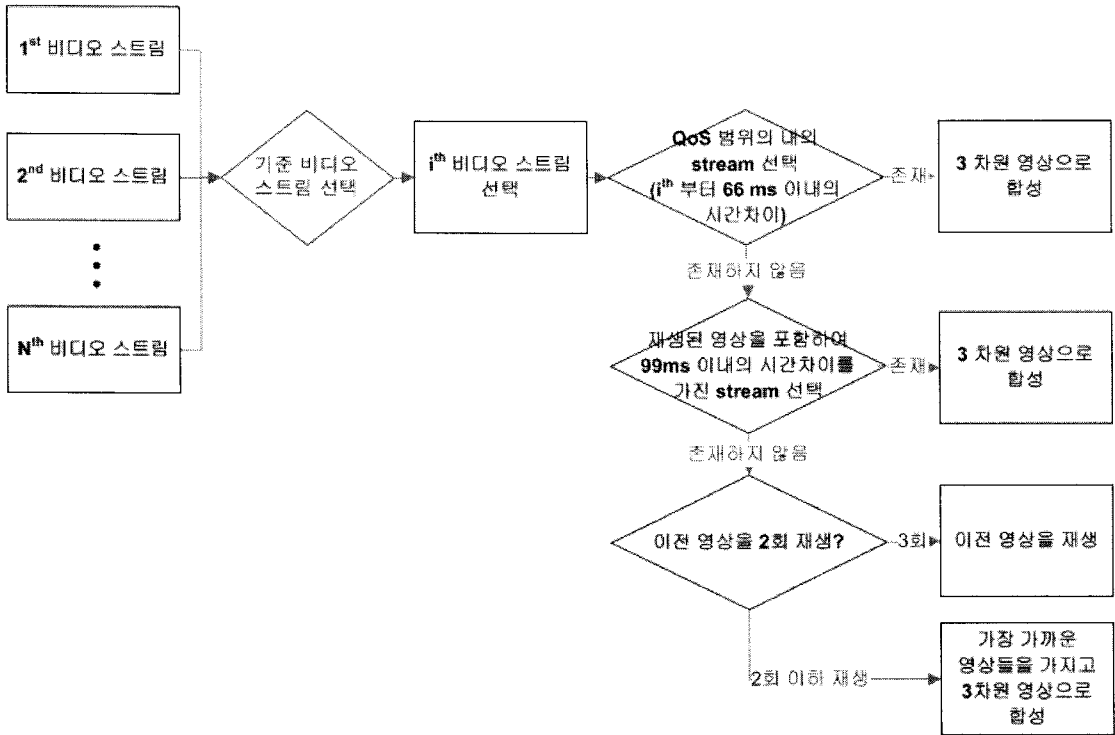


그림 5. QoS 보장범위를 이용한 다중 비디오 스트림 간의 동기화 과정

상을 연속하여 재생하는 횟수는 2회까지만 가능하도록 한다. 이미 2회 연속으로 재생된 경우, 기본 스트림의 시점을 제외한 다른 시점 다른 비디오 스트림들을 중에서 프레임간의 시간차이가 가장 적은 프레임들을 찾는다. 작은 프레임들을 이용하여 3차원 영상으로 합성하고, 3차원 디스플레이 장치에 영상이 재생한다. 이러한 기술을 이용하면, 영상을 재생하는 동안 자연스러운 움직임을 유지하면서 다중 비디오 스트림 간의 동기화를 제어할 수 있다. 이는 하나의 동영상을 구성하는 오디오와 비디오 간의 동기화만을 대상으로 하던 기존 방법으로는 해결할 수 없었던 3차원 동영상 동기화를 가능하게 한다.

### V. 구현 및 실험 결과

본 논문에서 DMOS 실험을 통해, 사람들에게 비디오 스트림을 3차원 영상으로 합성하여 입체감이 유지되는, 다중 비디오 스트림 시간차이 허용 오차 범위를 측정했고, 여기서 얻어낸 시간 차이 오차 허용범위를 동기화 제어 기술에 적용하여 스테레오 비디오의 동기화 제어를 수행하였다. 제안 시스템의 효율성을 평가하기 위하여 수행한 실험 조건을 표 1

에 나타냈다. 실험에 사용한 영상은 2장에서 DMOS 실험으로 사용했던 동기화된 스테레오 영상을 사용하였다. 본 실험에서는 데이터의 전송과 수신을 담당하는 컴퓨터를 한 대씩 사용하였다. 전송을 담당하는 컴퓨터에서는 미디어의 취득, 압축 및 패킷 전송을 수행하였고, 수신을 담당하는 컴퓨터에서는 패킷의 수신, 압축 복원 및 3차원 영상 재생 기능을 한다.

영상의 취득 및 재생은 표 1에 표시한 것과 같이 15fps로 설정했다. 네트워크 파라미터로는 지연은 50~150ms로 설정하고, 지터는 0~100ms로 설정하

표 1. 실험 조건

테스트 영상	Diving, Performance, Robot1, Soccer2, Street Organ, Trapeze
영상 해상도	QVGA (320×240)
비디오 코덱	H.264/AVC Baseline
영상 취득	15 fps
지연	0 ~ 150ms
버퍼크기	50ms
지터	0 ~ 100ms

었다. 이러한 네트워크의 파라미터는 WANem(Wide Area Network Emulator) 에뮬레이터를 사용하여 지연과 지터를 조절하였다. WANem은 TATA Consultancy Services Ltd. 에서 제공하는 네트워크 에뮬레이터 프로그램이다. 제안한 동기화 제어 기술을 적용한 실시간 다중 비디오 통신 시스템을 해당 조건의 네트워크에서 테스트하면서 네트워크 지연 시간을 제외한 순수 시스템 처리 시간 즉, 스테레오 영상을 취득하여 전송하고, 수신하여 합성 및 재생하는 순간까지의 시간을 측정하였다. 이는 제안한 알고리즘이 스테레오 통신 시스템을 제어하여 품질에 영향을 줄 수 있는 총 시간으로, 알고리즘이 유효하게 동작하는지를 확인할 수 있는 중요한 요소이다. 그림 6은 실험에 사용한 각 영상에 대한 시스템 처리 시

간의 측정 결과를 보여준다. 세로축은 시스템에서 걸린 시간 (단위 : ms)을 나타내고, 가로 축은 프레임 번호를 표시하였다. 구현한 실시간 다중 비디오 통신 시스템을 사용하여 3차원 영상을 재생하는 데까지 걸리는 시간은 순수 네트워크 지연을 제외하면 모두 100ms 이하임을 확인하였다. 시스템 처리 시간이 100ms 이하로 유지되면 사람이 주관적으로 받아들이는 데 문제가 없음을 II장에서 확인한 바 있다. 다만, 테스트 결과를 보면, 대부분의 경우 20~25ms 정도 수준의 지연 시간을 유지하지만 35ms를 넘는 지연 시간을 가지는 프레임이 주기적으로 나타나는데, 이는 테스트에 사용한 운영체제의 스케줄러가 주기적으로 동작해야 할 다른 프로세스의 존재 여부를 확인하는 과정으로 인해 나타나는 것으

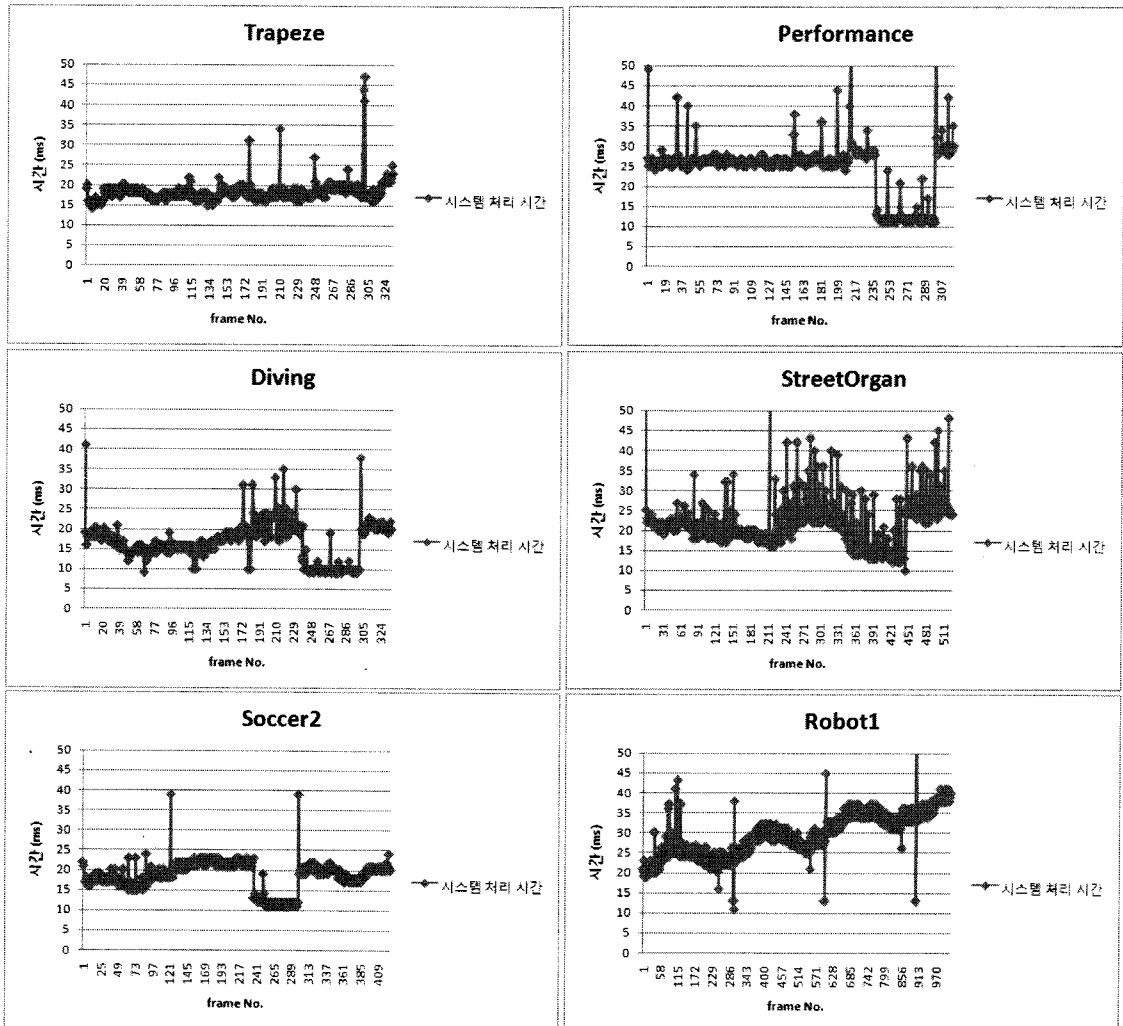


그림 6. 각 영상별 시스템 처리 시간



로 판단된다. 또한, 엄밀하게 보면 운영체제에 의한 지연은 순수한 시스템의 처리 시간으로 볼 수 없고, 최종적으로 사람이 느끼는데 있어서 문제가 있는 것으로 받아들이는 지연시간의 범위를 벗어나지도 않는다.

제한한 동기화 제어 기술의 성능을 검증하기 위해, 실시간 다중 비디오 통신 시스템에 제안 기술을 적용하여, 스테레오 비디오 동기화 및 합성으로 얻은 3차원 영상 결과로 주관적인 비디오품질 측정을 실시하여 DMOS를 얻었다. 본 실험에는 움직임이 적어 좌측과 우측 프레임간의 시간차이가 크게 발생하였을 때에도 3차원으로 인식하는 영상 등을 제외하고, 총 6개의 테스트 영상을 사용하여, 표 1의 실험 조건 하에 3차원 영상의 주관적 화질을 평가하였다. 두 컴퓨터 사이에 네트워크에서 발생할 지연을 50~150ms로 설정하여 지연에 대한 DMOS 값을 측정하였고, 지연을 100ms로 설정한 상태에서 지터를 0~100ms 까지 설정하여, 일정한 지연이 있는 상태에서 지터를 발생시킨 경우의 DMOS 값을 측정하였다. 테스트 결과를 그림 7에서 보여준다.

그림 7은 동시에 취득한 스테레오 영상을 기존의 동기화 제어 기술과 제안한 동기화 제어 기술을 적용한 실시간 다중 비디오 통신 시스템에 3차원 디스플레이 장치를 통해 재생하였을 때의 DMOS 측정 결과이다. 기존 결과는 지연 시간 허용 범위에 대한 제한 없이 동기화를 제어한 후에 생성한 영상을 보여준 결과에 대한 DMOS 값이다. 각 영상 별로 표시한 DMOS 값은 식 (11)에 나타낸 것과 같이, 네트워크 파라미터를 조절하면서 측정한 MOS 값과 원본영상을 측정한 MOS 값과의 차를 나타낸다. 그림 7에 표시한 그래프의 세로축은 DMOS의 값으로 0에 가까우면 좋은 영상이다. 가로 축은 ms 단위의 지연과 지터를 표시한 것으로써, 지연을 일반 숫자로 표시하고 지터는 괄호("()") 안에 표시하였다. 본 실험에서 DMOS 값이 5.1점에서 설명한 것과 같이 2.5 이하인 경우는 3차원 영상을 보여주었을 때에 자연스럽게 입체적으로 영상으로 인식한다는 것을 의미한다. 그림 7에서 보이는 것처럼, 지터가 버퍼보다 커서 손실이 발생한 경우에는 기존의 방법보다 제안한 기술이 더 자연스러운 3차원 영상을 만든다는 것을 보여주고 있다.

그림 7 (a)의 'Diving', (e)의 'Street Organ'을 제외한 나머지 영상들은 DMOS 값으로, 통계를 통해 여러 사람의 테스트 결과로 평균을 구하는 과정

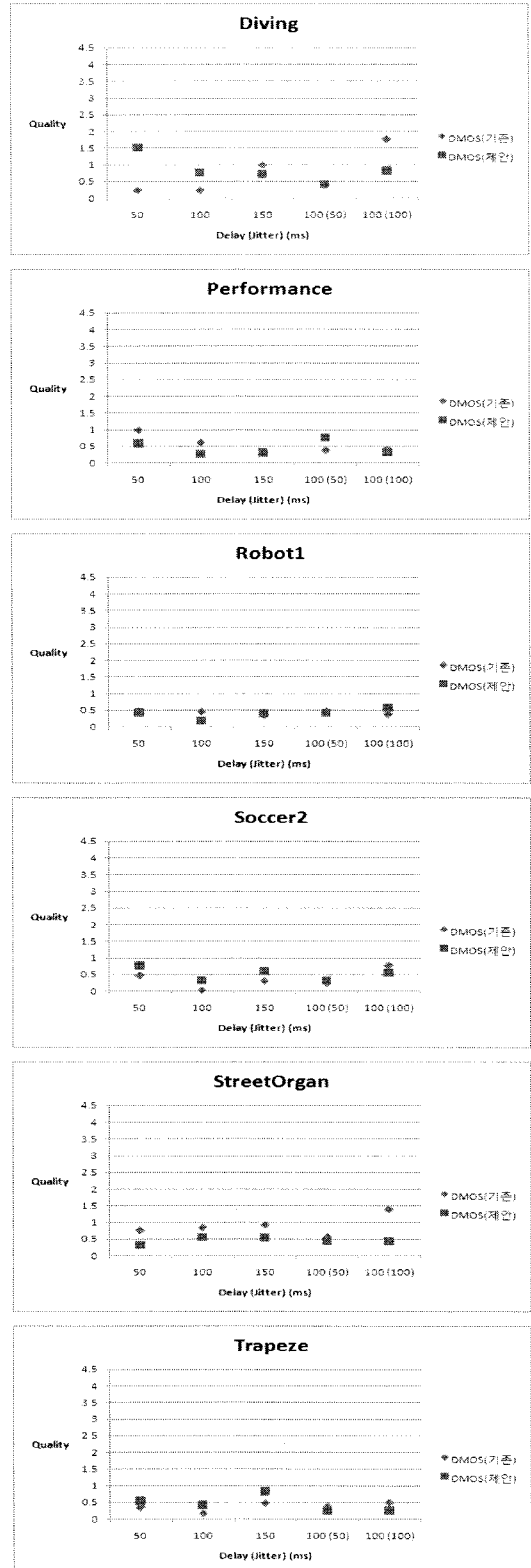


그림 7. 각 영상에 대한 DMOS 측정 값

에서 발생할 수 있는 표준편차 내에 존재하는 값들을 가진다. 이는 주어진 조건 하에서 송·수신함에 있어 해당 영상들이 3차원 영상으로 느끼는 데 지장이 없음을 의미한다. 'Diving'과 'Street Organ'의 경우 다른 영상들에 비해 압축 시 비트량이 많이 발생하기 때문에 네트워크의 영향을 많이 받는다. 할당된 버퍼의 범위를 넘어서는, 지터 150ms 이상의 구간에서 기존 방법의 DMOS 값이 0.8을 웃도는 경우가 많은 반면, 제안한 방법의 결과는 비교적 훌륭한 화질 구간인 0.8 이하의 DMOS 값을 가지는 결과를 확인할 수 있다. 이는 기존의 기술에 비해 일정한 네트워크에서 손실이 발생한 상황에서도 제안한 동기화 제어 기술은 손실에 대해서 시간차이 허용범위를 이용하여 다른 시점의 영상을 사용함으로써 기존의 방식에 비해 더욱 자연스럽게 3차원 영상으로 합성할 수 있다는 것을 의미한다.

## VI. 결 론

본 논문에서는 스테레오 비디오의 동기화를 위한 RTP/RTCP기반 동기화 방법을 제안하였다. MOS 실험을 통해 다중 비디오 스트림을 합성한 3차원 영상이 자연스럽게 입체적으로 느껴지는 뷰간 시간차이의 오차 허용범위를 측정했고, 이를 고려하여 스테레오 비디오 간의 동기화 제어 기술을 제안하였다. 제안한 기술을 실시간 다중 비디오 통신 시스템에 적용하여, 제안한 기술의 유효성을 검증하였다. 또한, 제안한 기술이 실제 네트워크 환경에서 적절하게 동작하는 것을 검증하기 위하여, 네트워크 파라미터를 다양하게 조절하여 주관적 비디오 서비스의 품질을 측정하는 실험을 하였다. 서비스 품질을 측정하는 DMOS 실험 결과, 제안하는 동기화 제어 기술은 기존의 방법에 비해 지터 등에 영향을 덜 받았고, 이를 통하여 합성한 3차원 영상은 DMOS 실험 스케일 중 매우 좋음과 좋음 범위에 속하는 것을 확인하였다.

## 참 고 문 헌

[1] 김은수, 이승현, 3차원 영상의 기초, 技多利, 1998.  
 [2] ISO/IEC MPEG & ITU-T VCEG, "Common test condition for multiview video coding," JVT-T207, Klagenfurt, Austria, July, 2006.  
 [3] ITU-R Recommendation BT.500-11, "Meth-

odology for the subjective assessment of the quality of television pictures," International Telecommunication Union, Geneva, Switzerland, Jan. 2002.

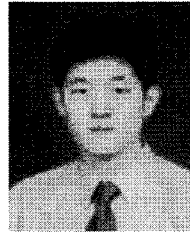
[4] I. Kouvelas, V. Hardman, and A. Watson, "Lip synchronization for use over the Internet: analysis and implementation," IEEE Global Telecommunications Conference, vol. 2, pp. 893-898, London, UK, Nov. 1996.  
 [5] K. H. Lee, D. H. Kim, M. G. Kang, K. H. Han, S. M. Park, and S. H. Kung, "An implementation of control protocol for multipoint audio-video teleconferencing systems," Proceedings of ICOIN-12. pp. 38-41, Tokyo, Japan, Jan. 1998.  
 [6] 김찬우, 박성준, 서광덕, "화상 전화 단말기에서의 효율적인 오디오/비디오 동기화 방법," 한국컴퓨터종합학술대회, vol. 32, no. 1(A), pp. 355-357, 2005.  
 [7] "RTP profile for audio and video conferences with minimal control," Internet Engineering Task Force, RFC 3551, July 2003.  
 [8] "RTP: A transport protocol for real-time applications," Internet Engineering Task Force, RFC3550, July. 2003.  
 [9] Ralf Steinmetz, "Human perception of jitter and media synchronization," IEEE Journal on Selected Areas in Communications, vol. 14, no. 1, Jan. 1996.  
 [10] H. Jinzenji and K. Hagishima, "Real-time audio and video broadcasting of IEEE GLOBECOM'96 over the Internet using new software," IEEE Communications Magazine, vol. 35, no. 4, pp. 34-38, Apr. 1997.  
 [11] Gerold Blakowski and Ralf Steinmetz, "A media synchronization survey : Reference model, specification, and case studies," IEEE Journal on Selected Areas in Communications, vol. 14, no. 1, Jan. 1996  
 [12] M. Narbutt and L. Murphy, "Adaptive playout buffering for audio/video transmission over the Internet," Proceedings of the 17th IEE UK Teletraffic Symposium, vol. 27, pp. 1-6, May 2001.  
 [13] Orion Hodson, Colin Perkins, and Vicky

Hardman, "Skew detection and compensation for Internet audio applications," Proceedings of the IEEE International Conference on Multimedia and Expo, vol. 3, pp. 1687-1690, New York, NY, July 2000.

- [14] J.-C. Bolot, "End-to-end packet delay and loss behavior in the Internet," Computer Comm. Review, vol. 23, no. 4 pp. 289-298, Sept. 1993.
- [15] 김병용, 이동진, 심동규, 권재철, "초저지연 비디오 통신을 위한 RTP 기반 립싱크 제어 기술에 관한 연구," 한국멀티미디어학회, vol. 10, no. 8, Aug. 2007.

이 동 진 (Dong-Jin Lee)

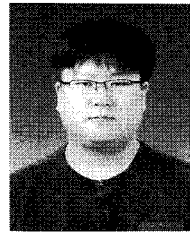
정회원



2006년 8월 광운대학교 컴퓨터 공학과 졸업  
2008년 8월 광운대학교 컴퓨터 공학과 석사  
2008년 8월~현재 NHN 근무  
<관심분야> 비디오 통신시스템

이 선 오 (Seon-Oh Lee)

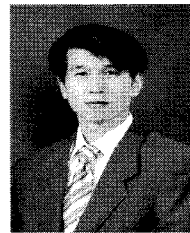
정회원



2007년 2월 광운대학교 컴퓨터 공학과 졸업  
2007년~현재 광운대학교 컴퓨터공학과 석박통합과정  
<관심분야> 영상품질, 멀티미디어 시스템

심 동 규 (Dong-Gyu Sim)

종신회원



1999년 서강대학교 전자공학과 공학박사  
1999년~2000년 (주)현대전자  
2000년~2002년 (주)바로비전  
2002년~2005년 University of Washington  
2005년~현재 광운대학교 컴퓨터공학과 교수

<관심분야> 영상신호처리, 멀티미디어 통신 시스템

이 혁 준 (Hyuk-Joon Lee)

종신회원



1993년 Syracuse University, Computer Science (박사)  
1994년~1996년 삼성전자(주) 멀티미디어 연구소 선임연구원  
1996년~현재 광운대학교 컴퓨터공학과 교수  
<관심분야> 무선 ad hoc 네트워크

위킹, 차량간 통신